Diffusion-weighted Imaging Is a Sensitive and Specific Magnetic Resonance Sequence in the Diagnosis of Ankylosing Spondylitis

Linda A. Bradbury, Kelly A. Hollis, Benoît Gautier, Sateesh Shankaranarayana, Philip C. Robinson, Nivene Saad, Kim-Anh Lê Cao, and Matthew A. Brown

ABSTRACT. ***Objective.*** We tested the discriminatory capacity of diffusion-weighted magnetic resonance imaging (DWI) and its potential as an objective measure of treatment response to tumor necrosis factor inhibition in ankylosing spondylitis (AS).

Methods. Three cohorts were studied prospectively: (1) 18 AS patients with Bath Ankylosing Spondylitis Disease Activity Index > 4, and erythrocyte sedimentation rate > 25 and/or C-reactive protein > 10 meeting the modified New York criteria for AS; (2) 20 cases of nonradiographic axial spondyloarthritis (nr-axSpA) as defined by the Assessment of Spondyloarthritis international Society (ASAS) criteria; and (3) 20 non-AS patients with chronic low back pain, aged between 18 and 45 years, who did not meet the imaging arm of the ASAS criteria for axSpA. Group 1 patients were studied prior to and following adalimumab treatment. Patients were assessed by DWI and conventional magnetic resonance imaging (MRI), and standard nonimaging measures.

Results. At baseline, in contrast to standard nonimaging measures, DWI apparent diffusion coefficient (ADC) values showed good discriminatory performance [area under the curve (AUC) > 80% for Group 1 or 2 compared with Group 3]. DWI ADC values were significantly lower posttreatment (0.45 ± 0.433 before, 0.154 ± 0.23 after, p = 0.0017), but had modest discriminating capacity comparing pre– and posttreatment measures (AUC = 68%). This performance was similar to the manual Spondyloarthritis Research Consortium of Canada (SPARCC) scoring system.

Conclusion. DWI is informative for diagnosis of AS and nr-axSpA, and has moderate utility in assessment of disease activity or treatment response, with performance similar to that of the SPARCC MRI score. (First Release February 15 2018; J Rheumatol 2018;45:771–8; doi:10.3899/jrheum.170312)

Key Indexing Terms:
ANKYLOSING SPONDYLITIS SPONDYLOARTHRITIS ADALIMUMAB
DIFFUSION MAGNETIC RESONANCE IMAGING SENSITIVITY SPECIFICITY

The development of magnetic resonance imaging (MRI) protocols for the diagnosis of axial spondyloarthritis (axSpA) has enabled earlier detection of disease in patients prior to the development of the radiographic abnormalities that define the development of ankylosing spondylitis (AS). This has revolutionized clinical care of these patients and research into the early phases of axSpA, raising the possibility that through early detection and intervention, the natural history of this disease may be altered and progressive spinal fusion avoided. While MRI has proven sensitive for early disease detection, reading MRI scans remains at least partially subjective, and significant variability in scoring remains problematic[1].

The paucity of data about MRI changes, particularly in spinal images, has raised questions about the specificity of currently used criteria for the diagnosis or assessment of axSpA by MRI[2,3]. Further, quantitative MRI measures of

From the Institute of Health and Biomedical Innovation, Queensland University of Technology, Translational Research Institute, Princess Alexandra Hospital; University of Queensland Diamantina Institute, Translational Research Institute, Princess Alexandra Hospital; Royal Brisbane and Women's Hospital; The University of Queensland, School of Medicine, Brisbane, Queensland, Australia.

The study was funded by a grant from Abbvie Australia. MAB is funded by an Australian National Health and Medical Research Council (NHMRC) Senior Principal Research Fellowship. KALC is supported by an NHMRC Career Development Fellowship (APP1087415). The funding bodies had no role in the design, collection, analysis, or interpretation of this study.

L.A. Bradbury, MSc, MNPSt, Institute of Health and Biomedical Innovation, School of Biomedical Sciences, Queensland University of Technology at Translational Research Institute; K.A. Hollis, BScN, RN, Institute of Health and Biomedical Innovation, School of Biomedical Sciences, Queensland University of Technology at Translational Research Institute; B. Gautier, PhD, University of Queensland Diamantina Institute, Translational Research Institute, Princess Alexandra Hospital; S. Shankaranarayana, MBBS, PhD, Princess Alexandra Hospital; P.C. Robinson, PhD, FRACP, Royal Brisbane and Women's Hospital; N. Saad, MD, FRANZCR, Princess Alexandra Hospital; K.A. Lê Cao, PhD, University of Queensland Diamantina Institute, Translational Research Institute, Princess Alexandra Hospital; M.A. Brown, MD, PhD, Institute of Health and Biomedical Innovation, School of Biomedical Sciences, Queensland University of Technology at Translational Research Institute.

Address correspondence to Prof. M.A. Brown, Institute of Health and Biomedical Innovation, Queensland University of Technology, Translational Research Institute, Princess Alexandra Hospital, Woolloongabba, Brisbane, Australia. E-mail: Matt.brown@qut.edu.au

Accepted for publication November 7, 2017.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2018. All rights reserved.

Downloaded on April 15, 2024 from www.jrheum.org

either inflammatory change or damage require complex scoring systems, and have therefore not become widely used in clinical practice, although they are used routinely in clinical trials where expert central reading is typical.

Diffusion-weighted imaging (DWI) is an MRI sequence that relies on the random Brownian motion of water molecules within tissues. This motion, or diffusion, is influenced by tissue composition and architecture, with reduced mean free water path being associated with greater signal. DWI has been shown to have advantages over standard MRI sequences in some clinical settings, including the early diagnosis of ischemic stroke and in staging some cancer types[4,5]. Different measures of diffusion have been proposed, with the apparent diffusion coefficient (ADC) measure the most widely used. Inflammation leads to higher ADC values through increased water in extracellular, less constrained, spaces. Several studies have investigated the clinical utility of DWI in AS, with suggestive evidence that this sequence has valuable discriminatory capacity between AS and noninflammatory back pain, and is sensitive to changes in disease activity in response to treatment with tumor necrosis factor-α inhibitors (TNFi)[6,7,8,9,10,11].

Disease activity measurement in AS is challenging. Currently the main approaches are to use patient self-completed questionnaires [such as the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)], blood tests [such as erythrocyte sedimentation rate (ESR)/C-reactive protein (CRP)], and combinations of questionnaires and blood tests [such as the Ankylosing Spondylitis Disease Activity Score (ASDAS)-CRP and ASDAS-ESR]. Patient self-completed questionnaires such as these are subjective and influenced by variation such as patient reporter biases, coexistent fibromyalgia, and concomitant medicines such as analgesics; they were not developed for use in assessing patients with noninflammatory back pain. In the current study we investigated the discriminatory capacity of DWI in comparison with standard classification criteria for axSpA, and assessed its usefulness as an objective measure of disease activity both cross-sectionally and longitudinally in response to treatment with the TNFi adalimumab, in comparison with standard clinical measures used for treatment response assessment.

MATERIALS AND METHODS

Subjects. The study protocol was approved by the Metro South Hospital and Health Service (approval HREC/11/QPAH/479) and The University of Queensland Research Ethics Committees (approval 20130000333), and all patients gave written informed consent. Three groups of patients were recruited and studied prospectively. Group 1 consisted of 18 patients diagnosed with AS defined by the modified New York criteria[12], all with active disease (BASDAI > 4, and CRP > 10 mg/dl and/or ESR > 25 mm/h). These patients underwent an MRI immediately prior to commencing 13 weeks of TNFi treatment (adalimumab 40 mg s/c fortnightly) and then again after 3 months of treatment. Group 2 consisted of 20 patients who met the Assessment of Spondyloarthritis international Society (ASAS) classification for nonradiographic axSpA (nr-axSpA)[13]. Group 3 consisted of 20 patients who had been referred to our AS clinic for diagnostic purposes with chronic low back pain for more than 3 months, and who did not meet the imaging arm of the ASAS criteria for axSpA. They were considered on clinical grounds (history and physical examination) not to have axSpA, although 10 of these patients did meet the ASAS clinical criteria for axSpA. In some analyses, Group 1 and 2 patients, classified as having inflammatory spinal arthritis, were pooled (active) and compared with Group 3 patients who were classified as having noninflammatory spinal arthritis (controls).

Excluded from participation were pregnant women, people under 18 years of age, people with an intellectual or mental impairment, and those with a contraindication for undergoing an MRI. Patients taking corticosteroids were also excluded, and nonsteroidal antiinflammatory drugs and analgesic medications were kept stable during the followup study of Group 1 patients. Standard exclusion criteria for use of TNFi therapies were applied to participants in Group 1.

For all groups, at each scan timepoint, disease activity was fully assessed by nonimaging methods — blood tests (full blood count, CRP, and ESR), patient-reported disease activity (BASDAI), functional score (Bath Ankylosing Spondylitis Functional Index; BASFI), mobility measure (Bath Ankylosing Spondylitis Metrology Index; BASMI), and the ASDAS (Ankylosing Spondylitis Disease Activity Score).

MRI methods. All patients underwent MRI scans of the entire spine and sacroiliac joints (SIJ) without contrast agent. The scans were performed on a 1.5-Tesla platform (Magnetom Avanto, Siemens) using up to 32 channels (spine array coil, 24 channels; and body array coil, 6 elements).

Three regions were scanned using the following protocols:

1. Cervicothoracic spine from the foramen magnum to superior border of T10 with sagittal fast spin-echo T1WI (TR/TE: 425/14) and sagittal short-tau inversion recovery (STIR; TR/TE: 4250/55), field of view (FOV) 38 cm, 3 mm thick, 20 slices; and
2. Thoracolumbar spine from T9 to S3 with sagittal fast spin-echo T1WI (TR/TE: 425/14) and sagittal STIR (TR/TE: 4250/55), FOV 38 cm, 4 mm thick, 20 slices; and
3. SIJ with (a) oblique coronal images parallel to the long axis of the sacrum: fast spin-echo T1WI (TR/TE: 521/14) and STIR (TR/TE: 5272/50), FOV 30 cm, 4 mm thick, 20 slices, (b) oblique axial images perpendicular to the long axis of the sacrum: fast spin-echo T1WI (TR/TE: 652/14) and STIR (TR/TE: 5140/61), FOV 24 cm, 4 mm thick, 0.4 mm intersection gap, 20 slices, and (c) oblique coronal images DWI single-shot echo planar sequence (TR/TE: 4100/87), b-value/diffusion gradients of 0, 50, 400, and 800s/mm^2, FOV 30 cm, 4 mm thick, 20 slices, 8 excitations.

Image interpretation and analysis. Nonblinded review of the scans was performed by a musculoskeletal radiologist (NS) on an Agfa Impax 6.1 picture archiving and communication system (PACS). Images were initially assessed for presence or absence of bone marrow edema (BMO) on the STIR sequence. The intensity of BMO was graded following visual inspection using: grade 0 = absent BMO, grade 1 = minor BMO less intense than adjacent vessels, grade 2 = marked BMO as intense as adjacent vessels. The Spondyloarthritis Research Consortium of Canada (SPARCC) MRI scores were determined according to protocol[14], several months after ADC mapping and BMO scoring.

ADC maps were linked to the corresponding DWI and STIR images using an image-linking tool on PACS to ensure accurate localization of the edema/region of interest (ROI). Because the ROI had to be reproduced on pre- and posttreatment scans, it was not possible to conduct a blinded analysis. The ADC value was measured twice, then averaged in both SIJ and mid-body of S2 vertebral level. The latter is termed "sacral ADC" in analyses (presumed normal reference bone). A circular ROI with a constant area of 0.44 cm^2 was placed on selected areas of BMO on the ADC map. Adjacent cortical bone, sclerosis, or fat metaplasia were excluded from the ROI to avoid volume averaging and erroneous ADC values. In patients where no BMO was detected, the ADC values were measured at identical points on both SIJ. Background reference signal was defined as the marrow signal along the midline of the sacrum at S2 level, which was also measured twice and averaged.

Statistical method. Between-group comparisons were performed using

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2018. All rights reserved.

Downloaded on April 15, 2024 from www.jrheum.org

Student t tests and ANOVA, with the exception of Group 1 pre– and posttreatment values, for which paired t tests/ANOVA were used. Correlation between values was assessed using either Pearson correlation coefficient (ADC values) or Spearman's correlation coefficient (clinical criteria). Discriminatory capacity was tested using receiver-operator characteristic (ROC) analysis. Nonparametric Mann-Whitney U tests (comparing independent groups) or Wilcoxon tests (comparing the paired pre– vs posttreatment in Group 1) assessed whether there was a significant difference between the ROC curves. AUC and thresholds for the best sensitivities/specificities were reported. Responsiveness was reported as the effect size using Glass's delta method, and standardized response mean (SRM) for the paired Group 1 pre- and posttreatment values. A significant difference was declared when $p < 0.05$.

RESULTS

Demographics and clinical values. Clinical and demographic details of the patients studied are provided in Table 1. Most disease activity measures were higher in Group 1 patients pretreatment than either Group 2 or 3 patients, including CRP, ESR, ASDAS-CRP, and ASDAS-ESR. BASDAI was higher in Group 1 than Group 2, but not higher than Group 3. BASFI values were also higher in Group 1 than Group 2 and 3, though not achieving statistical significance compared with Group 3 ($p = 0.078$). BASMI was not significantly different between the groups. Posttreatment, Group 1 patients had significantly lower disease activity values than pretreatment and either Group 2 or Group 3 patients. BASFI values fell significantly posttreatment, and were not different compared with either Group 2 or 3 patients. BASMI values did not change significantly with treatment. No significant differences were noted between Group 2 and 3 patients.

Strong correlation was noted between disease activity measures, both in the overall dataset ($r^2 > 0.3$), and considering active patients alone (Supplementary Figures 1 and 2, available with the online version of this article). Correlation between disease activity measures and BASFI was moderate (Spearman correlation coefficient 0.28–0.69).

Comparing the capacity of nonimaging measures to distinguish between active cases and controls, for most values no significant discriminatory capacity was found for BASDAI, BASFI, BASMI, ASDAS-ESR (AUC = 0.51–0.64, $p > 0.05$), with moderate discriminatory capacity observed for CRP and ESR alone (AUC = 0.74 and 0.69, and $p = 0.002$ and 0.02, respectively; Supplementary Figure 3, available with the online version of this article). For CRP, using a threshold of 6 to define active disease had a sensitivity of 80% and specificity of 68%.

Considering the discriminatory capacity of pretreatment measures in Group 1 patients in relation to posttreatment values, disease activity values all had high discriminatory capacity (AUC = 0.88–0.99, $p < 0.05$), with BASDAI and ASDAS-ESR and ASDAS-CRP having nearly equivalent high discriminatory capacity (AUC = 0.97–0.99, $p < 0.05$; Supplementary Figure 4, available with the online version of this article).

ADC values. DWI ADC values for each group are presented in Table 2. Because ADC values were similar between duplicate measures (Pearson correlations between reads were 0.976–0.993), these were averaged for further analyses. Of note, no Group 3 subjects displayed signs of inflammation by MRI.

Comparisons of ADC values at matched sites are presented in Table 3. Sacral ADC values were significantly higher in Group 1 patients pretreatment compared with Group 3 patients ($p = 0.0089$), but not different between any other groups ($p > 0.05$). Sacroiliac ADC values were significantly higher in Group 1 patients pretreatment and Group 2 patients compared with either Group 1 posttreatment or Group 3 patients. No difference was noted between Group 1 posttreatment and Group 3 patients, or between Group 1 pretreatment and Group 2 patients ($p > 0.05$).

For Group 1 pretreatment and Group 2, ADC values were higher over each SIJ than the reference sacral measurement ($p < 0.005$ all analyses; Supplementary Table 1, available with the online version of this article). In Group 1 posttreatment cases, this observation was less notable and significant only for the left but not right SIJ ($p = 0.0047$ and $p = 0.083$ respectively); no differences were observed in Group 3 patients.

Significant variation in ADC values was observed even among control subjects, with strong correlation between sacral and sacroiliac values within that Group (r^2 between sacral and left and right ADC values = 0.84 and 0.80,

Table 1. Demographic details of participants.

	Group 1		Group 2	Group 3
	Pretreatment	Posttreatment		
No.	18		20	20
Age, yrs, mean (range)	38 (23–68)		40.6 (23–62)	37 (20–60)
Sex (% male)	56		55	37
BASDAI	5.5	2	4.7	6.5
CRP	24.2	4.9	8.9	5.5
ESR	41.8	12	15.3	15.1
% HLA-B27 carriage	100		80	70

BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2018. All rights reserved.

Downloaded on April 15, 2024 from www.jrheum.org

Table 2. DWI ADC scores for different groups regarding each measurement in duplicate. Active group includes Group 1 pretreatment and Group 2. Duplicate measurements per side were averaged prior to the tests. Corrected values indicate sacroiliac sides subtracted from sacral and averaged.

	Group 1 Pretreatment		Group 1 Postreatment		Group 2		Active		Group 3	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
ADC.right.1	0.672	0.425	0.408	0.144	0.947	0.409	0.820	0.434	0.403	0.146
ADC.right.2	0.665	0.414	0.448	0.124	0.949	0.423	0.818	0.438	0.400	0.150
ADC.left.1	0.831	0.543	0.598	0.356	0.802	0.492	0.815	0.509	0.415	0.150
ADC.left.2	0.846	0.517	0.609	0.339	0.835	0.502	0.840	0.502	0.416	0.160
ADC.sacrum.1	0.301	0.148	0.357	0.143	0.330	0.183	0.317	0.166	0.438	0.137
ADC.sacrum.2	0.306	0.163	0.367	0.140	0.332	0.178	0.820	0.434	0.439	0.136
Corrected score	0.450	0.433	0.154	0.23	0.552	0.273	0.505	0.354	–0.03	0.07

DWI: diffusion-weighted imaging; ADC: apparent diffusion coefficient.

Table 3. Comparisons of sacral and sacroiliac ADC values comparing different groups at matched sites. P values from paired t tests (Group 1 pre- vs posttreatment) or t tests (any other comparison) are displayed. Duplicate measurements per side were averaged prior to the tests. Corrected values indicate sacroiliac sides subtracted from sacral and averaged. For example, on the left side, Group 1 pretreatment scores were significantly different compared with Group 1 posttreatment scores (p = 0.0028), whereas comparing with Group 2 scores at this site was marginally significant (p = 0.05).

	Group 1 Posttreatment				Group 2				Group 3			
	Left	Right	Sacral	Corrected	Left	Right	Sacral	Corrected	Left	Right	Sacral	Corrected
Group 1 pretreatment	0.0028	0.0074	0.079	0.0017	0.89	0.23	0.48	0.39	0.0011	0.0084	0.0089	0.00021
Group 1 posttreatment	—	—	—	—	0.050	0.0012	0.85	1.7×10^{-5}	0.025	0.26	0.080	0.0040
Group 2					—	—	—	—	0.00015	9.0×10^{-5}	0.085	2.0×10^{-9}

ADC: apparent diffusion coefficient.

respectively), suggesting that variation in underlying tissue properties between individuals is present. Therefore, for further analyses, right and left sacroiliac values were corrected for sacral values by subtracting sacral values from the sacroiliac values, which were then averaged to produce a single ADC value (corrected) for each individual (Table 2).

Comparing groups, significant differences were observed between corrected values for Group 1 pretreatment and posttreatment, and between Group 3 and any of the groups. A 2.9-fold reduction in mean corrected scores between pre– and posttreatment groups was observed [mean corrected ADC score Group 1 pretreatment 0.45 (SD 0.433), posttreatment 0.154 (SD 0.23), p = 0.0017]. While no difference was observed between Group 1 pretreatment and Group 2 values, a borderline significant reduction in Group 1 posttreatment values was noted compared with Group 2 values.

BMO and SPARCC scores correlate with ADC values. ADC values were substantially higher for increasing BMO scores (Supplementary Figure 5, available with the online version of this article). No Group 3 subject had a BMO score > 0, whereas 58% (11/19) of Group 1 subjects and 71% (15/21) of Group 2 subjects had a BMO score > 0. While BMO scores did decline in the Group 1 posttreatment versus pretreatment comparison, this was not significant on either side (left side p = 0.0024, right side p = 0.068).

Pretreatment SPARCC scores correlated well with ADC values (Group 1 pretreatment Spearman correlation coefficient = 0.79, p = 0.00010; Group 2 Spearman correlation coefficient = 0.86, p = 0.00001). Correlations were weaker posttreatment in Group 1 (Spearman correlation coefficient = 0.43, p = 0.073).

Correlation between ADC values and clinical disease activity measures. No consistent correlations were noted between right and left sacroiliac ADC values and any clinical disease activity measures. While nominally significant correlations for some measures were observed with individual sides, in each case these were either unilateral or with opposite directions of correlation with either side.

Discriminatory capacity and responsiveness of ADC measures. Excellent discriminatory capacity (AUC > 0.8, p < 0.05) was noted comparing Group 1 pretreatment, Group 2, and active cases with Group 3 controls (Figure 1; and Supplementary Figure 6, available with the online version of this article). Comparing active cases with controls, the optimal ADC threshold (0.054) resulted in a specificity of 95% and sensitivity of 87% (AUC = 0.91, p = 1.5×10^{-8}).

In contrast, no significant discrimination was observed in comparisons of Group 1 pretreatment and Group 2 cases (AUC = 0.45, p = 0.65). Moderate though statistically significant discriminatory capacity was observed comparing Group 1 pre– and posttreatment values (AUC = 0.67, p = 2.8 ×

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2018. All rights reserved.

Downloaded on April 15, 2024 from www.jrheum.org

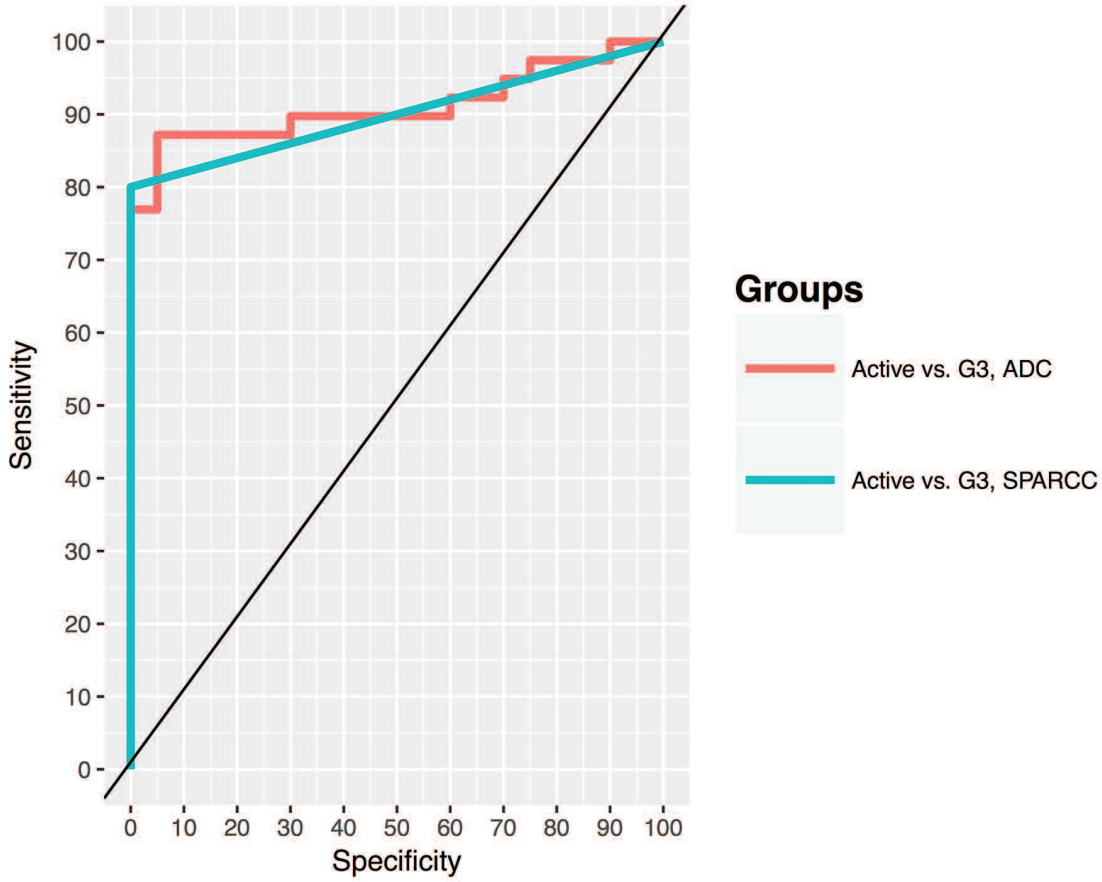

Figure 1. The discriminatory capacity of DWI in active cases and back pain controls was similar to the manual SPARCC scoring system. Active cases comprise Group 1 pretreatment and Group 2, and the controls by Group 3. ROC curve analysis for ADC DWI (AUC of 91%; $p = 1.5 \times 10^{-8}$) and SPARCC scores (AUC of 90%; $p = 1.9 \times 10^{-7}$). DWI: diffusion-weighted imaging; SPARCC: Spondyloarthritis Research Consortium of Canada; ROC: receiver-operator characteristic; ADC: apparent diffusion coefficient; AUC: area under the curve.

10^{-3}). For an ADC value threshold of 0.53, the test had a specificity of 0.94 and sensitivity of 0.56.

SPARCC scores performed similarly to ADC measures in these analyses, with AUC and p values respectively for Group 1 pretreatment, Group 2 and active cases versus Group 3 controls ≥ 0.8 and < 0.05. No difference was noted in Group 1 pretreatment and Group 2 cases (AUC = 0.52, p = 0.88). Moderate though statistically significant discriminatory capacity was observed comparing Group 1 pre– and posttreatment values (AUC = 0.7, $p = 2.5 \times 10^{-3}$).

ADC measure and SPARCC score effect sizes were similar when comparing active cases with Group 3 controls, and far greater than for all other measures (Table 4). In assessment of change in response to treatment (Group 1 pre- vs posttreatment), again ADC measures performed similarly to SPARCC scores, with similar effect sizes and SRM. In contrast to the cross-sectional comparison of active cases with Group 3 controls, in this longitudinal comparison both MRI scores showed much less responsiveness than any other score except metrology (BASMI), with which they performed similarly (Table 5).

DISCUSSION

Our study confirms that DWI has excellent capacity to distinguish active AS cases and nr-axSpA cases from control cases with noninflammatory back pain. Its performance exceeded that of widely used measures of disease activity (including BASDAI, ESR, CRP, and ASDAS-ESR and ASDAS-CRP) in distinguishing active disease from noninflammatory back pain controls. In our current study, the semiautomated DWI ADC scoring method performed similarly to manual SPARCC scoring in both baseline comparisons and in comparison of pre– with posttreatment imaging.

Our study demonstrated that following TNFi treatment, DWI ADC values at a group level did decrease significantly, with posttreatment values being 66% lower than pretreatment values. This is comparable with other quantitative MRI measures, such as the SPARCC sacroiliac score, in which the

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2018. All rights reserved.

Downloaded on April 15, 2024 from www.jrheum.org

Table 4. Responsiveness analysis of ADC scores in comparison of active group with Group 3. For SPARCC scores, because the SD for Group 3 is zero, no effect size can be calculated.

	Active Mean	Active SD	Group 3 Mean	Group 3 SD	Effect Size	p
ADC	0.50	0.35	–0.03	0.07	7.610	10^{-10}
SPARCC	10.60	12.06	0.00	0.00	NA	9.6×10^{-6}
ASDAS-CRP	3.42	1.28	2.77	0.89	0.736	0.049
ASDAS-ESR	3.37	1.25	2.88	0.94	0.529	0.14
ESR	27.84	23.08	15.10	13.98	0.911	0.026
CRP	16.13	19.00	5.51	6.57	1.618	0.0092
BASDAI	6.25	2.52	6.54	2.14	–0.135	0.78
BASFI	4.08	2.20	4.13	2.92	–0.019	0.94
BASMI	2.33	1.81	2.11	1.96	0.112	0.78

ADC: apparent diffusion coefficient; SPARCC: Spondyloarthritis Research Consortium of Canada; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; ASDAS: Ankylosing Spondylitis Disease Activity Score; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; BASFI: Bath Ankylosing Spondylitis Functional Index; BASMI: Bath Ankylosing Spondylitis Metrology Index.

Table 5. Responsiveness measures of magnetic resonance imaging scores and other measures in comparison of Group 1 cases pre- and posttreatment.

	Group 1 Pretreatment		Group 1 Posttreatment				
	Mean	SD	Mean	SD	Effect Size	SRM	p
ADC	0.45	0.43	0.15	0.23	–0.684	–0.877	0.0017
SPARCC	13.40	15.02	3.25	5.66	–0.676	–0.911	0.00083
ASDAS-CRP	4.32	0.62	1.75	0.99	–4.118	–2.273	7.9×10^{-8}
ASDAS-ESR	4.29	0.69	1.72	0.99	–3.701	–2.372	6.3×10^{-8}
ESR	41.78	24.03	12.00	12.22	–1.239	–1.335	6.3×10^{-5}
CRP	24.18	18.07	4.88	4.57	–1.068	–1.050	0.00063
BASDAI	7.94	1.13	2.01	1.72	–5.237	–3.396	5×10^{-10}
BASFI	5.02	1.45	2.70	2.02	–1.599	–1.029	0.00063
BASMI	2.78	2.14	1.82	2.00	–0.446	–0.953	0.00095

SRM: standardized response mean; ADC: apparent diffusion coefficient; SPARCC: Spondyloarthritis Research Consortium of Canada; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; ASDAS: Ankylosing Spondylitis Disease Activity Score; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; BASFI: Bath Ankylosing Spondylitis Functional Index; BASMI: Bath Ankylosing Spondylitis Metrology Index.

effect size and SRM in our study was very similar to those for ADC values. In the EMBARK study of etanercept (ETN) treatment of nr-axSpA, SPARCC sacroiliac scores fell at Week 12 by 48% compared with baseline[15] and by Week 48 to 58% of baseline values[16]. In a further study of ETN treatment in axSpA (radiographic status not stated), Berlin MRI scores fell at Week 24 by 60% compared with baseline, and by 69% at Week 48[17].

However, considering discriminatory capacity at an individual level, ADC measures had only moderate capacity to distinguish cases pre– and posttreatment (AUC = 0.68), and the clinical disease activity measures studied performed significantly better (AUC 0.88–0.95). The performance of SPARCC scoring was similar, and both scores showed lower responsiveness (as assessed by effect size or SRM) than all other scores except BASMI. There is a significant need for methods to identify axSpA cases that respond well to TNF inhibition. ESR, CRP, and MRI have previously been shown to have utility in this setting, largely in posthoc analysis of clinical trial data. It is possible that DWI may perform better after a longer period of treatment, and further trials are warranted.

DWI also did not correlate well in cross-sectional analyses with disease activity measures, either patient self-reported (BASDAI), acute-phase reactants (ESR or CRP), or combined measures (ASDAS-CRP, ASDAS-ESR). DWI images of the spine have a very limited spatial resolution and thus do not allow accurate localization or ADC measurement of active disease. This in turn precludes evaluation of patients with predominant active spinal disease. MRI has previously also been shown to have only minor correlations with disease activity measures. In established AS, neither the Berlin sacroiliac nor spinal scores correlated with BASDAI, morning stiffness, global pain, patient's or physician's global scores, CRP, ESR, or BASFI[18]. The Berlin sacroiliac score was also shown not to correlate with baseline patient global

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2018. All rights reserved.

Downloaded on April 15, 2024 from www.jrheum.org

scores, ESR, or CRP in AS, and to have only modest correlations with ESR ($\rho = 0.31$, $p = 0.016$) and CRP ($\rho = 0.38$, $p = 0.004$), and not with patient global scores, in undifferentiated SpA[19]. Similarly, in a study of patients with SpA (radiographic status unknown), baseline Berlin sacroiliac scores did not correlate with ASDAS, BASDAI, or CRP levels, and change in score with treatment did not correlate with change in BASDAI or CRP[20]. In studies of TNF inhibition in axSpA cases (of uncertain radiographic status), reduction in MRI scores was shown to correlate with reduction in BASDAI ($\rho = 0.37$) and CRP ($\rho = 0.52$ among those with an elevated CRP), although this correlation was only observed in those with short disease duration (< 4 yrs)[21]. Performance of the SPARCC score is similar, with no correlation noted between SPARCC scores and self-reported measures (variously BASDAI, nocturnal pain, patient's global assessment, and total back pain)[22,23,24], with 1 study showing correlation with change in SPARCC score and change in CRP with treatment[24]. Thus, at least cross-sectionally, current MRI quantitative scores do not consistently correlate with either self-reported or objective (CRP, ESR) measures of axSpA activity.

While we did not assess interreader reliability in this study, in 3 previous studies of DWI this has been good to excellent ($\kappa = 0.62$–0.93[10], ICC = 0.89–0.98[25], $\kappa = 0.89$[26]). Intrareader reliability in the current study was also excellent, with correlation between duplicate readings ranging from 0.976 to 0.993. These compare favorably with reported reliability statistics for Berlin and SPARCC MRI scores (intra- and interreader ICC, respectively, 0.93–0.97 and 0.5–0.97 Berlin score, 0.92–0.97 and 0.78–0.98 SPARCC score[27]).

Different approaches for quantifying DWI measures have been developed, including the ADC and intravoxel incoherent motion (IVIM), which may improve the performance of the methodology[28]. DWI reflects the random motion of water protons, which is performed using a range of diffusion weightings (b values), usually between 0 and 1000 s/mm^2. Water diffusion, tissue interfaces, membranes, and microcapillary perfusion influence on *in vivo* DWI. The ADC value is a monoexponential fit that represents the relationship between the log (tissue signal intensity/attenuation) and the b value. Signal attenuation at low b values ($\leq$ 100 s/mm^2) is strongly influenced by microcapillary perfusion, while signal attenuation at higher b values (100–1000 s/mm^2) reflects the diffusion component. IVIM is a more complex technique based on this concept, which separately quantifies microcapillary perfusion and pure tissue diffusion coefficient from DWI data, allowing more accurate characterization of the biexponential behavior of tissues as a result of microcapillary perfusion. Software required for the IVIM analysis is not widely available on all MRI platforms, and for that reason was not performed in our study. However, it has theoretical advantages over the ADC measurements, and comparisons of the 2 approaches in axSpA would be valuable.

This study shows that DWI has excellent performance as a diagnostic tool in distinguishing axSpA from noninflammatory causes of back pain, but performs only modestly in detecting treatment response at 13 weeks and correlates poorly with standard disease-activity measures. Its performance is very similar to the manual SPARCC scoring system. Further studies are indicated to investigate the performance of DWI in screening back pain patients for axSpA, and in distinguishing the subset of patients with nr-axSpA who go on to develop AS.

ACKNOWLEDGMENT

We thank the patients who took part in this study.

ONLINE SUPPLEMENT

Supplementary material accompanies the online version of this article.

REFERENCES

1. van Gaalen FA, Bakker PA, de Hooge M, Schoones JW, van der Heijde D. Assessment of sacroiliitis by radiographs and MRI: where are we now? Curr Opin Rheumatol 2014;26:384-8.
2. Weber U, Zhao Z, Rufibach K, Zubler V, Lambert RG, Chan SM, et al. Diagnostic utility of candidate definitions for demonstrating axial spondyloarthritis on magnetic resonance imaging of the spine. Arthritis Rheumatol 2015;67:924-33.
3. Deodhar A. Sacroiliac joint magnetic resonance imaging in the diagnosis of axial spondyloarthritis: "a tiny bit of white on two consecutive slices" may be objective, but not specific. Arthritis Rheumatol 2016;68:775-8.
4. Albers GW. Diffusion-weighted MRI for evaluation of acute stroke. Neurology 1998;51:S47-9.
5. Benz MR, Vargas HA, Sala E. Functional MR imaging techniques in oncology in the era of personalized medicine. Magn Reson Imaging Clin N Am 2016;24:1-10.
6. Zhao YH, Li SL, Liu ZY, Chen X, Zhao XC, Hu SY, et al. Detection of active sacroiliitis with ankylosing spondylitis through intravoxel incoherent motion diffusion-weighted MR imaging. Eur Radiol 2015;25:2754-63.
7. Pan C, Hu DY, Zhang W, Li XM. Role of diffusion-weighted imaging in early ankylosing spondylitis. Chin Med J (Engl) 2013;126:668-73.
8. Sanal HT, Yilmaz S, Kalyoncu U, Cinar M, Simsek I, Erdem H, et al. Value of DWI in visual assessment of activity of sacroiliitis in longstanding ankylosing spondylitis patients. Skeletal Radiol 2013;42:289-93.
9. Ai F, Ai T, Li X, Hu D, Zhang W, Morelli JN. Value of diffusion-weighted magnetic resonance imaging in early diagnosis of ankylosing spondylitis. Rheumatol Int 2012;32:4005-13.
10. Bozgeyik Z, Ozgocmen S, Kocakoc E. Role of diffusion-weighted MRI in the detection of early active sacroiliitis. AJR Am J Roentgenol 2008;191:980-6.
11. Gaspersic N, Sersa I, Jevtic V, Tomsic M, Praprotnik S. Monitoring ankylosing spondylitis therapy by dynamic contrast-enhanced and diffusion-weighted magnetic resonance imaging. Skeletal Radiol 2008;37:123-31.
12. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
13. Rudwaleit M, Landewe R, van der Heijde D, Listing J, Brandt J, Braun J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part I): classification of paper patients by expert

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2018. All rights reserved.

Downloaded on April 15, 2024 from www.jrheum.org

opinion including uncertainty appraisal. Ann Rheum Dis 2009;68:770-6.
14. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Williams M, Stone M, et al. Spondyloarthritis research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:703-9.
15. Dougados M, van der Heijde D, Sieper J, Braun J, Maksymowych WP, Citera G, et al. Symptomatic efficacy of etanercept and its effects on objective signs of inflammation in early nonradiographic axial spondyloarthritis: a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheumatol 2014;66:2091-102.
16. Maksymowych WP, Dougados M, van der Heijde D, Sieper J, Braun J, Citera G, et al. Clinical and MRI responses to etanercept in early non-radiographic axial spondyloarthritis: 48-week results from the EMBARK study. Ann Rheum Dis 2016;75:1328-35.
17. Song IH, Hermann K, Haibel H, Althoff CE, Listing J, Burmester G, et al. Effects of etanercept versus sulfasalazine in early axial spondyloarthritis on active inflammatory lesions as detected by whole-body MRI (ESTHER): a 48-week randomised controlled trial. Ann Rheum Dis 2011;70:590-6.
18. Rudwaleit M, Schwarzlose S, Hilgert ES, Listing J, Braun J, Sieper J. MRI in predicting a major clinical response to anti-tumour necrosis factor treatment in ankylosing spondylitis. Ann Rheum Dis 2008;67:1276-81.
19. Xu M, Lin Z, Deng X, Li L, Wei Y, Liao Z, et al. The Ankylosing Spondylitis Disease Activity Score is a highly discriminatory measure of disease activity and efficacy following tumour necrosis factor-alpha inhibitor therapies in ankylosing spondylitis and undifferentiated spondyloarthropathies in China. Rheumatology 2011;50:1466-72.
20. Pedersen SJ, Sorensen IJ, Hermann KG, Madsen OR, Tvede N, Hansen MS, et al. Responsiveness of the Ankylosing Spondylitis Disease Activity Score (ASDAS) and clinical and MRI measures of disease activity in a 1-year follow-up study of patients with axial spondyloarthritis treated with tumour necrosis factor alpha inhibitors. Ann Rheum Dis 2010;69:1065-71.
21. Weiss A, Song IH, Haibel H, Listing J, Sieper J. Good correlation between changes in objective and subjective signs of inflammation in patients with short- but not long duration of axial spondyloarthritis treated with tumor necrosis factor-blockers. Arthritis Res Ther 2014;16:R35.
22. Lambert RG, Salonen D, Rahman P, Inman RD, Wong RL, Einstein SG, et al. Adalimumab significantly reduces both spinal and sacroiliac joint inflammation in patients with ankylosing spondylitis: a multicenter, randomized, double-blind, placebo-controlled study. Arthritis Rheum 2007;56:4005-14.
23. Maksymowych WP, Dhillon SS, Park R, Salonen D, Inman RD, Lambert RG. Validation of the spondyloarthritis research consortium of Canada magnetic resonance imaging spinal inflammation index: is it necessary to score the entire spine? Arthritis Rheum 2007;57:501-7.
24. Maksymowych WP, Salonen D, Inman RD, Rahman P, Lambert RG, Group CS. Low-dose infliximab (3 mg/kg) significantly reduces spinal inflammation on magnetic resonance imaging in patients with ankylosing spondylitis: a randomized placebo-controlled study. J Rheumatol 2010;37:1728-34.
25. Dallaudiere B, Dautry R, Preux PM, Perozziello A, Lincot J, Schouman-Claeys E, et al. Comparison of apparent diffusion coefficient in spondylarthritis axial active inflammatory lesions and type 1 Modic changes. Eur J Radiol 2014;83:366-70.
26. Boy FN, Kayhan A, Karakas HM, Unlu-Ozkan F, Silte D, Aktas I. The role of multi-parametric MR imaging in the detection of early inflammatory sacroiliitis according to ASAS criteria. Eur J Radiol 2014;83:989-96.
27. Hededal P, Østergaard M, Sorensen IJ, Loft AG, Hindrup JS, Thamsborg G, et al. Development and validation of MRI sacroiliac joint scoring methods for the semiaxial scan plane corresponding to the Berlin and SPARCC MRI scoring methods, and of a new global MRI sacroiliac joint method. J Rheumatol 2018;45:70-7.
28. Koh DM, Collins DJ, Orton MR. Intravoxel incoherent motion in body diffusion-weighted MRI: reality and challenges. AJR Am J Roentgenol 2011;196:1351-61.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2018. All rights reserved.

Downloaded on April 15, 2024 from www.jrheum.org